

SOFICITY FOR MONOIDS, SEMIGROUPS, AND GENERAL DYNAMICAL SYSTEMS

JAN CANNIZZO

ABSTRACT. We examine several definitions of soficity for monoids obtained by generalizing various definitions of sofic groups. They are not all equivalent and include the definition recently introduced by Ceccherini-Silberstein and Coornaert. One of these definitions readily generalizes to semigroups (albeit in an arguably unsatisfying way), thus addressing a question asked by Kambites. We conclude by proposing a definition of soficity for a general dynamical system consisting of a semigroup acting by measure-preserving transformations on a probability space.

1. INTRODUCTION

Since its introduction, the notion of soficity has proved to be of fundamental interest in several branches of mathematics. *Sofic groups* were first defined by Gromov [11], who showed that they satisfy Gottschalk’s surjunctivity conjecture in topological dynamics, and they were given their name by Weiss [14] shortly thereafter (the word “sofic” is derived from the Hebrew word for “finite”). Sofic groups are closely linked to a number of other conjectures, notably Connes’ embedding conjecture and the determinant conjecture in the theory of von Neumann algebras (see [9] and the survey of Pestov [13] for more), and it was a breakthrough when Bowen [3] defined a notion of *sofic entropy* for measure-preserving actions of sofic groups, thereby extending the classical entropy theory developed for \mathbb{Z} -actions by Kolmogorov and Sinai.

It was soon realized that the notion of soficity can be extended to objects other than groups. By formulating soficity in terms of the weak convergence of measures (so-called *Benjamini-Schramm convergence*, introduced in [2]), it became possible to speak of *sofic random graphs*—see [1]—and, mutatis mutandis, *sofic random Schreier graphs* (see, for instance, [5] for a definition). Using this idea, Elek and Lippner [8] went on to define the more general notion of a *sofic equivalence relation*, and working in greater generality still, Dykema, Kerr, and Pichot [7] have recently defined *sofic groupoids* (see also [4]).

In each of the aforementioned contexts, the following question—arguably the most naïve question it is possible to ask—has resisted proof and remains open as of this writing.

Question 1.1. *Does there exist a nonsofic object?*

Although the general opinion seems to be that nonsofic objects should exist, finding an example of such an object has proven to be difficult. It was therefore remarkable when Ceccherini-Silberstein and Coornaert [6] recently defined *sofic monoids* and went on to answer Question 1.1 by showing that there does exist a nonsofic monoid. In fact

their example is not a particularly exotic object: it is the well-known *bicyclic monoid*, namely the monoid with presentation

$$B = \langle a, b \mid ab = e \rangle,$$

where $e \in B$ denotes the identity. Further work on sofic monoids was carried out by Kambites [12] who, building on [6], exhibited a large class of sofic monoids.

Yet Ceccherini-Silberstein and Coornaert’s definition of a sofic monoid, although certainly natural from an algebraic point of view, can be said to mark a departure from other definitions of soficity. Soficity as it is defined for groups, random graphs, equivalence relations, and groupoids is very much connected with ergodic theory—in particular, the presence of an invariant measure—whereas this connection appears to be absent from Ceccherini-Silberstein and Coornaert’s theory. In fact it is automatic in the theories of other sofic objects that the existence of an underlying invariant measure is a prerequisite for defining soficity in the first place; hence, it is natural to ask whether it is possible to develop a theory of sofic monoids that takes invariance into account.

As noted by Kambites, soficity as defined in [6] “imposes no restriction whatsoever on the internal complexity of the monoid outside the group of units” ([12], p. 12). Indeed, Proposition 4.7 of [6] implies that, though the bicyclic monoid B is nonsofic, the monoid $\hat{B} := B \cup \{e'\}$ obtained by adjoining a new element e' to B and defining e' to be the identity is sofic—a striking fact, given that the internal structure of \hat{B} is identical to that of B outside of the point e' . Kambites concludes by remarking that

If seeking applications in semigroup theory more widely, one is drawn to ask if there is an alternative, probably stronger, definition of a sofic monoid which also generalises sofic groups but exerts more control on the internal structure of the rest of the monoid. A natural test of whether a definition is satisfactory in this respect would be whether the resulting class is closed under the taking of monoid subsemigroups.... Such a definition, if found, is also likely to extend naturally to semigroups without an identity element ([12], p. 12).

One of the aims of this paper is to explore what such a definition might look like, beginning with the observation that a sofic group (just like an *amenable group*) may naturally be defined in terms of *approximately invariant measures*. We do, in fact, arrive at an alternative definition of a sofic monoid which readily generalizes to semigroups, but a semigroup is sofic in this new sense if and only if it embeds into a sofic group—arguably a heavy-handed condition. On the other hand, the definition is natural in that it provides a clear connection with ergodic theory, and although many semigroups (including the bicyclic monoid) remain “nonsofic” according to this definition, they are nonsofic for obvious reasons—roughly speaking, because they are not *invariant structures*. This same phenomenon appears in the theories of other sofic objects as well: Given a probability measure on the space of Schreier graphs of a finitely generated group which is not conjugation-invariant, the corresponding random Schreier graph might be said to be “nonsofic,” but this is rather uninteresting, as it is a consequence of the definition of a sofic random Schreier graph that its law is conjugation-invariant. The principle here is that one must first restrict one’s attention to invariant structures, and only then ask whether they are sofic.

To conclude, we offer another way to relate semigroups and soficity, namely by passing from semigroups to their actions on spaces equipped with an invariant measure. We thus formulate a notion of a general *sofic dynamical system*. This generalizes the notion of a sofic equivalence relation introduced by Elek and Lippner and may be an interesting direction for future research.

This paper is organized as follows: In Section 2, we review background material and single out four equivalent definitions of sofic groups, each of which generalizes in a natural way to monoids. In Section 3, we show how these various generalizations—one of which is the definition of Ceccherini-Silberstein and Coornaert—relate to one another and, in particular, are not all equivalent. Finally, in Section 4, we introduce the notion of a general sofic dynamical system, provide a basic example of such a system, and pose several questions related to our definition.

2. SOFIC GROUPS AND SOFIC MONOIDS

Sofic groups are, speaking very roughly, groups which admit finite approximations. They are a common generalization of amenable and residually finite groups (although there exist sofic groups which are neither amenable nor residually finite) and, as we will see, can be defined in a number of equivalent ways. We first establish some preliminaries.

If S is a semigroup and \mathcal{A} is a generating set for S , then the (*left*) *Cayley graph* $\Gamma = \Gamma(S, \mathcal{A})$ of S constructed with respect to \mathcal{A} is the graph whose vertex set is identified with S and whose edge set consists of all triples (s, a, as) , where $a \in \mathcal{A}$ is a generator and $s \in S$ is an arbitrary element. The edge (s, a, as) is understood to be directed from s to as , which are its initial and terminal points, respectively, and labeled with a . The distance $d_{\mathcal{A}}(s, t)$ between two vertices $s, t \in \Gamma$ is defined to be the length of a shortest sequence of edges, or path,

$$(x_0, a_0, y_0), \dots, (x_n, a_n, y_n)$$

such that $x_0 = s$, $y_n = t$, and $y_i = x_{i+1}$ for all $0 \leq i \leq n-1$, provided that such a path exists. We set $d_{\mathcal{A}}(s, s) = 0$ for all $s \in \Gamma$. Note that $d_{\mathcal{A}}$ is in general not a metric, since in a general semigroup, $d_{\mathcal{A}}(s, t)$ need not be defined (in fact it is possible for Γ to be disconnected, e.g. if S is a free semigroup of rank $n > 1$).

Given a vertex $s \in \Gamma$, we define the ball of radius $r \in \mathbb{N}$ centered at s , denoted $B_r(\Gamma, s)$, or just $B_r(s)$ for short, to be the subgraph of Γ induced by the vertices at distance less than or equal to r from s . If S is a group and \mathcal{A} is symmetric, so that $a \in \mathcal{A}$ implies $a^{-1} \in \mathcal{A}$, then $B_r(s)$ coincides with the ball of radius r taken with respect to the usual graph metric on Γ , but as before, note that this fails to be true for a general semigroup.

If S has an identity e , then it is natural to regard $\Gamma(S, \mathcal{A})$ as a *rooted graph* (Γ, e) by distinguishing the vertex e . Given two rooted graphs (Γ, x) and (Δ, y) whose edges are directed and labeled with elements of a set \mathcal{A} (we do not assume here that Γ and Δ are Cayley graphs), we define a graph homomorphism $\phi : \Gamma \rightarrow \Delta$ to be a graph homomorphism in the usual sense which also satisfies $\phi(x) = y$ and maps one edge to another in such a way that respects the direction and labeling. A graph homomorphism is a graph isomorphism if it is invertible.

If X is a set, we denote by $\text{End}(X)$ the semigroup of all endomorphisms (that is, self-maps) $f : X \rightarrow X$, where the binary operation is composition of functions. If X is finite, recall that the (*normalized*) *Hamming metric* on $\text{End}(X)$ is defined to be

$$d_{\text{Ham}}(f, g) := \frac{1}{|X|} |\{x \in X \mid f(x) \neq g(x)\}|.$$

Given a semigroup S , we now define a (K, ε) -*action* of S on a finite set as follows (see also [9] and [13]).

Definition 2.1. Let S be a semigroup, $\varepsilon > 0$ a real number, and $K \subseteq S$ a finite subset of S . A (K, ε) -*action* of S on a finite set X is a function $\psi : K \rightarrow \text{End}(X)$ such that the following hold:

- i. If $s, t, st \in K$, then $d_{\text{Ham}}(\psi(st), \psi(s) \circ \psi(t)) \leq \varepsilon$.
- ii. If S has an identity e and $e \in K$, then $d_{\text{Ham}}(\psi(e), \text{Id}_X) \leq \varepsilon$, where Id_X is the identity map on X .
- iii. For all distinct $s, t \in K$, one has $d_{\text{Ham}}(\psi(s), \psi(t)) \geq 1 - \varepsilon$.

It can be useful to think of a (K, ε) -action as a finite model of a (free) semigroup action in which a certain amount of error is allowed. We are now in a position to define sofic groups and will in fact present four equivalent definitions.

Definition 2.2. The following are equivalent definitions of a sofic group.

- (1) A group G with finite generating set \mathcal{A} is sofic if for any $r \geq 0$ and any $\varepsilon > 0$, there exists a finite \mathcal{A} -labeled graph Γ such that

$$\frac{|\{x \in \Gamma \mid B_r(x) \cong G_{r, \mathcal{A}}\}|}{|\Gamma|} \geq 1 - \varepsilon,$$

where $B_r(x)$ is the ball of radius r centered at $x \in \Gamma$ and $G_{r, \mathcal{A}}$ is the ball of radius r centered at the identity in the Cayley graph $\Gamma(G, \mathcal{A})$.

- (2) A group G is sofic if for any finite subset $K \subseteq G$ and any $\varepsilon > 0$, it admits a (K, ε) -action $\psi : K \rightarrow \text{End}(X)$ on a finite set X .
- (3) A group G is sofic if for any finite subset $K \subseteq G$ and any $\varepsilon > 0$, it admits a (K, ε) -action $\psi : K \rightarrow \text{End}(X)$ on a finite set X such that the uniform probability measure μ_X on X is approximately invariant, in the sense that

$$\|\mu_X - \psi(g)_* \mu_X\| \leq \varepsilon$$

for all $g \in K$, where $\|\cdot\|$ denotes the total variation norm.

- (4) A group G is sofic if it acts essentially freely on a measure space (X, μ) equipped with an invariant, finitely additive probability measure μ defined on the set of all subsets of X .

Definition (1) is the original definition of Gromov. Note that, although it applies only to finitely generated groups, soficity may be regarded as a local property: by Definitions (2) or (3), an arbitrary group is sofic if and only if all of its finitely generated subgroups are sofic. The condition of being approximable by finite labeled graphs is called the

Weiss condition in [6]. Definition (2) is due to Elek and Szabó. Definition (3) is a generalization of a definition of an amenable group: A group G is *amenable* if for any finite subset $K \subseteq G$ and any $\varepsilon > 0$, there exists a finite subset $X \subseteq G$ such that μ_X is approximately invariant, in the sense that

$$\|\mu_X - g_*\mu_X\| \leq \varepsilon$$

for all $g \in K$. Finally, Definition (4) is again due to Elek and Szabó [9]. It too generalizes a known definition of an amenable group: A finitely generated group G is amenable if and only if it admits an invariant, finitely additive probability measure defined on the set of all subsets of G .

Definitions (1)-(4) do not represent an exhaustive list of definitions of a sofic group, but they are such that any one of them readily generalizes to a definition of a sofic monoid. Accordingly, we formulate the following definition.

Definition 2.3. We say that a monoid is (k) -sofic if it is sofic according to the definition obtained by replacing the word “group” with the word “monoid” in part (k) of Definition 2.2 above.

A monoid is sofic in the sense of Ceccherini-Silberstein and Coornaert if it is (2)-sofic, and one might guess that, given that they are equivalent for groups, Definitions (1)-(4) will turn out to be equivalent for monoids as well. This is not the case, however. We will show that (1)-soficity and (2)-soficity are indeed equivalent, thereby clarifying an issue raised by Ceccherini-Silberstein and Coornaert, but that (1)-soficity and (2)-soficity are not equivalent to (3)-soficity, which implies (4)-soficity.

We do not wish to dwell on the question of which of these definitions, if any, yields the “right” notion of soficity for monoids. We feel that this is a matter of taste, and it is not our aim to tout one definition as right and another as wrong. We do wish, however, to explore these varying definitions in some detail, and to point out how (3)-soficity and (4)-soficity align more naturally with soficity as it is defined for other structures.

3. RELATING DIFFERENT NOTIONS OF SOFCITY FOR MONOIDS

Ceccherini-Silberstein and Coornaert show (see Theorem 6.1 of [6]) that (1)-soficity implies (2)-soficity for monoids, but they establish the converse only under the assumption that the monoid is left-cancellative. Our first result is that equivalence holds in full generality.

Theorem 3.1. *A monoid is (1)-sofic if and only if it is (2)-sofic.*

Proof. As mentioned above, it is established in [6] that (1)-soficity implies (2)-soficity. Conversely, let M be a finitely generated (2)-sofic monoid, \mathcal{A} a finite generating set of M , and $M_{r,\mathcal{A}}$ the r -neighborhood of the identity in the Cayley graph of M constructed with respect to \mathcal{A} .

Observe that it is sufficient to treat the case when $r = 1$, since it is always possible to pass to a larger generating set. That is, constructing a graph that locally looks like

$M_{r,\mathcal{A}}$ is tantamount to constructing a graph that locally looks like $M_{1,\mathcal{B}}$, where \mathcal{B} is, say, the generating set consisting of all monoid elements in $M_{r,\mathcal{A}}$.

Accordingly, let K consist of all monoid elements in $M_{1,\mathcal{A}}$, and let $\psi : K \rightarrow \text{End}(X)$ be a (K, ε) -action of M on a finite set X . Note that $\mathcal{A} \subseteq K$. For simplicity, write $\psi_s := \psi(s)$, and endow X with a graph structure by taking as its set of edges all triples of the form $(x, s, \psi_s(x))$, where $s \in K$. The edge $(x, s, \psi_s(x))$ is understood to be directed from x to $\psi_s(x)$ and labeled with s . Denote by μ_X the uniform probability measure on X .

By Proposition 6.2 of [6], we may assume that $\psi_e = \text{Id}_X$. For distinct elements $s, t \in K$, define the subset $A_{s,t} \subseteq X$ as

$$A_{s,t} := \{x \in X \mid \psi_s(x) = \psi_t(x)\}.$$

For elements $s, t \in K$ such that $st \in K$, define the subset $A_{s,t,st} \subseteq X$ as

$$A_{s,t,st} := \{x \in X \mid (\psi_s \circ \psi_t)(x) \neq \psi_{st}(x)\}.$$

By the definition of a (K, ε) -action, each of the sets $A_{s,t}$ and $A_{s,t,st}$ has measure less than or equal to ε , so that

$$\mu_X \left(\bigcup_{s \neq t} A_{s,t} \right) \leq |K|^2 \varepsilon \quad (3.1)$$

and

$$\mu_X \left(\bigcup_{s,t,st \in K} A_{s,t,st} \right) \leq |K|^2 \varepsilon =: \frac{\delta}{2}. \quad (3.2)$$

Denoting by A the union of all sets $A_{s,t}$ and $A_{s,t,st}$, it follows from the bounds (3.1) and (3.2) that $\mu_X(X \setminus A) \geq 1 - \delta$. Clearly, $1 - \delta$ can be made arbitrarily close to 1 by choosing ε to be sufficiently close to 0.

Given a point $x \in X \setminus A$, we would like to exhibit an isomorphism $\phi : M_{1,\mathcal{A}} \rightarrow B_1(x)$ such that $\phi(e) = x$, where $B_1(x)$ denotes the 1-neighborhood of x . If $x \in X \setminus A$, then consider the graph homomorphism $\phi : M_{1,\mathcal{A}} \rightarrow B_1(x)$ given by

$$\phi(s, t, ts) = (\psi_s(x), t, (\psi_t \circ \psi_s)(x)) = (\psi_s(x), t, \psi_{ts}(x)).$$

Let (s, t, ts) and $(s', t', t's')$ be two edges in $M_{1,\mathcal{A}}$. By construction, if $s \neq s'$, then $\psi_s(x) \neq \psi_{s'}(x)$. Likewise, if $ts \neq t's'$, then $\psi_{ts}(x) \neq \psi_{t's'}(x)$, which shows that ϕ is an embedding. It remains to show that ϕ is in fact surjective. To this end, note that it is impossible that there exist an edge $(x, s, \psi_s(x))$ in $B_1(x)$ which is not in the image of ϕ , since for every $x \in X$, there is exactly one s -labeled outgoing edge attached to x and it follows from the definition of ϕ that this edge is $\phi(e, s, s)$. Suppose $(y, s, \psi_s(y))$, where $y \neq x$, is some other edge in $B_1(x)$ which is not in the image of ϕ . Since y belongs to the 1-neighborhood of x , there exists some $t \in K$ such that

$$(x, t, \psi_t(x)) = (x, t, y)$$

is an edge in $B_1(x)$. But then

$$(x, st, \psi_{st}(x)) = (x, st, (\psi_s \circ \psi_t)(x)) = (x, st, \psi_s(y))$$

is an edge in $B_1(x)$ as well. The definition of ϕ and the fact that y has a unique s -labeled outgoing edge attached to it now imply that $(y, s, \psi_s(y))$ must have belonged to the image of ϕ after all. \square

Our next goal is to demonstrate that (3)-soficity for monoids is a very strong condition: a monoid is (3)-sofic if and only if it embeds into a group. Before proving this claim, we establish a lemma which asserts, roughly speaking, that under certain conditions a labeled graph which always looks the same when looking forwards must always look the same when looking backwards. To be more precise, if Γ is a graph whose edges are directed and labeled with elements of an alphabet \mathcal{A} (an \mathcal{A} -labeled graph), denote by $\bar{\Gamma}$ the graph obtained by keeping all edge labels of Γ the same but reversing the direction of each edge. Given a vertex $x \in \Gamma$, let

$$B_{\infty}^{+}(\Gamma, x) = \varinjlim B_r(\Gamma, x),$$

and let

$$B_{\infty}^{-}(\Gamma, x) = B_{\infty}(\bar{\Gamma}, x).$$

As before, if there is no confusion about the graph we are referring to, we may write $B_{\infty}^{+}(x)$ instead of $B_{\infty}^{+}(\Gamma, x)$, etc. Note that if Γ is the Cayley graph (or, more generally, a Schreier graph) of a group constructed with respect to a symmetric generating set \mathcal{A} , then $\Gamma = \bar{\Gamma}$.

Lemma 3.2. *Let Γ be a connected \mathcal{A} -labeled graph such that for each $x \in \Gamma$ and each $a \in \mathcal{A}$, there exists precisely one outgoing and one incoming edge labeled with a attached to x , and such that for any two vertices $x, x' \in \Gamma$, the neighborhoods $B_{\infty}^{+}(x)$ and $B_{\infty}^{+}(x')$ are isomorphic. Then Γ is the Cayley graph of a group.*

Proof. Let $x, x' \in \Gamma$ be two vertices. By assumption, there exists an isomorphism $\phi : B_{\infty}^{+}(x) \rightarrow B_{\infty}^{+}(x')$ of rooted labeled graphs. If $y \in B_{\infty}^{+}(x)$, put $y' := \phi(y)$. We claim that ϕ extends to an isomorphism

$$\phi : B_{\infty}^{+}(x) \cup B_{\infty}^{-}(y) \rightarrow B_{\infty}^{+}(x') \cup B_{\infty}^{-}(y').$$

Suppose it did not. Then there would exist a vertex $z \in B_{\infty}^{-}(y)$, obtained by starting at y and following a word w in the alphabet \mathcal{A}^{-1} , together with distinct paths γ_1 and γ_2 which begin at z and terminate at the same point, whereas the point $z' \in B_{\infty}^{-}(y')$, reached by starting at y' and following the word w , would be such that the corresponding paths γ'_1 and γ'_2 did not terminate at the same point. Either this, or the same situation would occur with the roles of $B_{\infty}^{-}(y)$ and $B_{\infty}^{-}(y')$ reversed. In either case, it is obvious that $B_{\infty}^{+}(z)$ and $B_{\infty}^{+}(z')$ would not be isomorphic, a contradiction. It follows by induction that ϕ extends to an isomorphism

$$\phi : \bigcup_{y \in B_{\infty}^{+}(x)} B_{\infty}^{-}(y) \rightarrow \bigcup_{y \in B_{\infty}^{+}(x')} B_{\infty}^{-}(y)$$

between the full backwards orbits of $B_{\infty}^{+}(x)$ and $B_{\infty}^{+}(x')$.

Suppose next that $x_0 \in B_{\infty}^{+}(x)$ is a *deficient* vertex, i.e. a vertex such that $y := a^{-1}(x_0) \notin B_{\infty}^{+}(x)$ for some $a \in \mathcal{A}$. Put $y' := a^{-1}(\phi(x_0))$. We claim that ϕ also extends to an isomorphism

$$\phi : B_{\infty}^{+}(x) \cup B_{\infty}^{+}(y) \rightarrow B_{\infty}^{+}(x') \cup B_{\infty}^{+}(y').$$

If it did not, then there would exist a word w in the alphabet \mathcal{A} such that the path obtained by beginning at y and following w terminates at a vertex $z \in B_{\infty}^{+}(x) \cap B_{\infty}^{+}(y)$ but the path obtained by beginning at y' and following w terminates at a vertex $t' \notin B_{\infty}^{+}(x')$. Either this, or the same situation would occur with the roles of $B_{\infty}^{+}(y)$ and

$B_\infty^+(y')$ reversed. Now let w' be a word in the alphabet \mathcal{A} such that the path obtained by starting at y' and following w terminates at $z' := \phi(z)$ (since $z' \in B_\infty^+(y')$, such a path exists). We must then conclude that $B_\infty^+(y)$ and $B_\infty^+(y')$ are not isomorphic: The paths obtained by beginning at y and following w and w' , respectively, terminate at the same point z , whereas the corresponding paths beginning at y' terminate at distinct points $t' \neq z'$. This is a contradiction.

It follows, finally, that by iterating the previous arguments, ϕ can be extended to a full isomorphism $\phi : (\Gamma, x) \rightarrow (\Gamma, x')$. One can first extend ϕ to an isomorphism between the backwards orbits of $B_\infty^+(x)$ and $B_\infty^+(x')$, then adjoin the forward orbits of a deficient vertex and extend ϕ further, and so on. But a connected \mathcal{A} -labeled graph with the property that any two choices of root yield isomorphic rooted labeled graphs is necessarily the Cayley graph of a group. \square

The next theorem shows that (3)-soficity is a very strong condition.

Theorem 3.3. *A monoid is (3)-sofic if and only if it embeds into a sofic group.*

Proof. Let M be a (3)-sofic monoid and $\psi : K \rightarrow \text{End}(X)$ a (K, ε) -action of M on a finite set X such that $\|\mu_X - (\psi_s)_*\mu_X\| \leq \varepsilon$ for all $s \in K$ (we again set $\psi_s := \psi(s)$ for notational convenience). Suppose that, for a given $s \in K$, there exists a set $A \subseteq X$ of measure $\mu_X(A) > \varepsilon$ such that $|\psi_s^{-1}(x)| > 1$ for all $x \in A$. Then

$$\begin{aligned} \|\mu_X - (\psi_s)_*\mu_X\| &= \max_{E \subseteq X} |\mu_X(E) - ((\psi_s)_*\mu_X)(E)| \\ &= \max_{E \subseteq X} |\mu_X(E) - \mu_X(\psi_s^{-1}(E))| \\ &\geq |\mu_X(A) - \mu_X(\psi_s^{-1}(A))| \\ &\geq |\mu_X(A) - 2\mu_X(A)| \\ &= \mu_X(A) > \varepsilon, \end{aligned}$$

a contradiction. By the same argument, if there exists a set $A \subseteq X$ of measure $\mu_X(A) > \varepsilon$ such that $|\psi_s^{-1}(x)| < 1$ for all $x \in A$, then we would have $\mu_X(\psi_s^{-1}(A)) = \mu_X(\emptyset) = 0$, again allowing us to deduce that $\|\mu_X - (\psi_s)_*\mu_X\| > \varepsilon$. It follows that for each $s \in K$, $|\psi_s^{-1}(x)| = 1$ for all x over a set of measure at least $1 - \varepsilon$, i.e. each ψ_s is approximately invertible.

Accordingly, there exists a subset $X_0 \subseteq X$ of measure at least $1 - |K|\varepsilon$ (which we can of course choose to be arbitrarily close to 1 by choosing ε to be sufficiently small) such that all ψ_s , where $s \in K$, satisfy $|\psi_s^{-1}(x)| = 1$ for any $x \in X_0$. Now endow X with the labeled graph structure whose edges consist of all triples of the form $(x, a, (\psi_n)_a(x))$, where $a \in \mathcal{A} \subseteq K$ belongs to a generating set of M . By Theorem 3.1 (note that (3)-soficity immediately implies (2)-soficity), there exists a subset $X_1 \subseteq X$ of measure arbitrarily close to 1 such that for each $x \in X_1$, the r -neighborhood $B_r^+(x)$ is isomorphic to $M_{r, \mathcal{A}}$. Thus, for all $x \in X_0 \cap X_1$, which again has measure arbitrarily close to 1, $B_r^+(x) \cong M_{r, \mathcal{A}}$ and x has exactly one outgoing and one incoming edge labeled with s attached to it for all $s \in K$.

Let $\{\psi_n : K_n \rightarrow \text{End}(X_n)\}_{n \in \mathbb{N}}$ be a sequence of (K_n, ε_n) -actions such that each K_n and each X_n is finite,

$$K_1 \subseteq \dots \subseteq K_n \subseteq \dots,$$

$\bigcup_{n \in \mathbb{N}} K_n = M$, and $\varepsilon_n \rightarrow 0$. We also choose our sequence in such a way that each of the uniform probability measures μ_n is approximately invariant. By the preceding argument, μ_n is not only approximately invariant with respect to each $(\psi_n)_s \in \text{End}(X_n)$, where $s \in K_s$, but approximately invariant with respect to each $(\psi_n)_s^{-1}$. The weak limit μ of the measures μ_n is therefore invariant with respect to the free group generated by \mathcal{A} . Moreover, μ is concentrated on rooted graphs (Γ, x) such that $B_\infty^+(x) \cong M_{\infty, \mathcal{A}}$ and such that x has exactly one incoming edge labeled with a attached to it for each $a \in \mathcal{A}$. It follows from Lemma 3.2 that μ is in fact concentrated on the Cayley graph of a group G , which is obviously a group into which M embeds. \square

Corollary 3.4. *The notion of (1)-soficity (or (2)-soficity) is not equivalent to (3)-soficity for monoids.*

Proof. Let M be a monoid which does not embed into a group. By Proposition 4.7 of [6], the monoid $\hat{M} := M \cup \{e'\}$ obtained by adjoining a new element e' to M and defining e' to be the identity is (2)-sofic and hence (1)-sofic. Since M does not embed into a group, \hat{M} does not embed into a group either. Therefore, \hat{M} is not (3)-sofic. \square

Corollary 3.5. *If a monoid is (3)-sofic, then it is (4)-sofic.*

Proof. Let M be a (3)-sofic monoid. Then M embeds into a sofic group G , which, by virtue of being sofic, admits an essentially free action $G \curvearrowright (X, \mu)$ on a measure space equipped with a finitely additive probability measure μ defined on the set of all subsets of X . The restriction of this action to M yields the desired essentially free action of M . \square

Remark 3.6. *We do not know whether (4)-soficity also implies (3)-soficity but conjecture a positive answer.*

A semigroup S need not have an identity. As such, carrying the notion of (1)-soficity (or, equivalently, (2)-soficity) over to semigroups is problematic: there is no canonical point in the Cayley graph of S whose neighborhood is to serve as a model for approximation. In light of Theorem 3.3, however, (3)-soficity can be applied to semigroups without an identity element and is indeed closed under the taking of monoid subsemigroups. The result is arguably an unsatisfying definition, as soficity for semigroups then turns out to be no more general a notion than soficity for groups: if a group G is sofic, then any subsemigroup of G is sofic, and, conversely, if a semigroup is sofic, then it embeds into a sofic group. But here we reiterate our point that (3)-soficity, like notions of soficity for other objects, retains a clear connection with ergodic theory.

4. A NOTION OF SOFICITY FOR GENERAL SEMIGROUP ACTIONS

Although we have found only a rather restrictive definition of soficity for semigroups, it is natural to pass from semigroups to their actions, whereupon a new idea presents

itself. This idea has already played out in the context of groups, where Elek and Lippner developed the notion of a sofic discrete measured equivalence relation [8]. This notion might equally well be called a *sofic dynamical system*, meaning a countable group acting by measure-preserving automorphisms on a Lebesgue space (recall that a *Lebesgue space*—also called a *standard probability space*—is a probability space whose nonatomic part is isomorphic to the unit interval equipped with Lebesgue measure). We will, in turn, consider a more general kind of dynamical system wherein groups are replaced by semigroups.

Definition 4.1. By a *dynamical system*, we will mean a countable semigroup S acting by measure-preserving endomorphisms on a Lebesgue space (X, μ) . We will work only with left actions (it is of course possible to develop an analogous theory for right actions) and also require that the action of our semigroup be *finite-to-one*, meaning that for any $s \in S$ and almost every $x \in X$, the preimage $s^{-1}(x)$ is finite.

We would like to define the notion of a sofic approximation to a system (X, μ, S) . In doing so, it is helpful to be able to visualize the action of the semigroup S . We do this in the following way. Given a point $x \in X$ and an element $s \in S$, we define the *preimage graph* $\Gamma_x(s)$ of x with respect to $s \in S$ to be the Schreier graph of the action of s restricted to the set $\{x\} \cup s^{-1}(x)$. Thus, $\Gamma_x(s)$ is the graph whose vertex set is $\{x\} \cup s^{-1}(x)$ and whose edge set consists of all triples (y, s, x) , where $y \in s^{-1}(x)$. We define a *partial Schreier graph* to be a connected union of preimage graphs. As the set of partial Schreier graphs is naturally ordered by inclusion, we define the *full Schreier graph* of x , denoted Γ_x , to be the maximal partial Schreier graph containing x .

Definition 4.2. We denote by $\text{Sch}(S, X)$ the *space of full Schreier graphs* of the action of S on X , taking each graph $\Gamma_x \in \text{Sch}(S, X)$ to be rooted at the vertex x .

Note that if S is a group, then a given Γ_x is the usual Schreier graph of S acting on the orbit of the point x , but a full Schreier graph is in general a significantly larger object, taking into account, as it were, many different forward and backward orbits.

Now let X be a standard Borel space. For the sake of concreteness, we will always assume that $X = \{0, 1\}^\infty$. The space X has a natural projective structure

$$X = \varprojlim X_r,$$

where $X_r = \{0, 1\}^r$ and the connecting maps $\pi_r : X_r \rightarrow X_{r-1}$ restrict a binary string of length r to its first $r - 1$ digits. An element $\omega \in X_r$ thus determines the *cylinder set*

$$C_\omega := \{x \in X \mid \pi_{\infty, r}(x) = \omega\},$$

where $\pi_{\infty, r} : X \rightarrow X_r$ is the natural projection onto X_r , and we denote by \mathfrak{A} the corresponding Borel σ -algebra on X . Given a finite-to-one action of a countable semigroup S on X , we may pass to the associated space of full Schreier graphs $\text{Sch}(S, X)$, which we endow with a projective structure as follows. Let $\{S_i\}_{i \in \mathbb{N}}$ be an enumeration of S . Define $\text{Sch}_r(S, X)$ to be the set of (isomorphism classes of) r -neighborhoods of the roots of full Schreier graphs $\Gamma_x \in \text{Sch}(S, X)$ which are spanned by the elements s_1, \dots, s_r (all other edges are neglected) and whose vertex-labels are truncated to their first r digits. It is easy to see that

$$\text{Sch}(S, X) = \varprojlim \text{Sch}_r(S, X),$$

where the connecting maps $\rho_r : \text{Sch}_r(S, X) \rightarrow \text{Sch}_{r-1}(S, X)$ are the obvious restriction functions. As before, an element $U \in \text{Sch}_r(S, X)$ determines the cylinder set

$$C_U := \{\Gamma_x \in \text{Sch}(S, X) \mid \rho_{\infty, r}(\Gamma_x) = U\},$$

where $\rho_{\infty, r} : \text{Sch}(S, X) \rightarrow \text{Sch}_r(S, X)$ is the natural projection onto $\text{Sch}_r(S, X)$, and we denote by \mathfrak{B} the corresponding Borel σ -algebra on $\text{Sch}(S, X)$. It should be pointed out that an element of $\text{Sch}_r(S, X)$ is to be understood as an r -neighborhood taken with respect to the usual graph metrics on full Schreier graphs Γ_x .

Our idea is to pass from X to $\text{Sch}(S, X)$, a space which records all of the information about the action of S . The map $f : (X, \mathfrak{A}) \rightarrow (\text{Sch}(S, X), \mathfrak{B})$ given by $f(x) = \Gamma_x$ is clearly a bijection, and it is easy to see that f^{-1} is measurable: Given a cylinder set $C_\omega \subseteq X$, where $\omega \in X_r$, let $\{U_i\}_{i \in I}$ be the collection of all partial Schreier graphs $U_i \in \text{Sch}_r(S, X)$ whose roots are labeled with ω . Then

$$f(C_\omega) = \bigcup_{i \in I} C_{U_i},$$

which is measurable (note that I is at most countable, since the action of S is finite-to-one). Let \mathfrak{A}' denote the completion of \mathfrak{A} with respect to the measure μ . Then $f_*^{-1}\mathfrak{B} \subseteq \mathfrak{A}'$ (the image of a Borel set under a measurable function is analytic and hence Lebesgue measurable), which allows us to push forward the measure μ , obtaining a measure $\nu := f_*\mu$ on $\text{Sch}(S, X)$. Let \mathfrak{B}' denote the completion of \mathfrak{B} with respect to ν .

Proposition 4.3. *The map $f : (X, \mathfrak{A}', \mu) \rightarrow (\text{Sch}(S, X), \mathfrak{B}', \nu)$ given by $f(x) = \Gamma_x$ is an S -equivariant isomorphism of Lebesgue spaces.*

Proof. It is clear that f is an S -equivariant bijection, since $s(\Gamma_x) = \Gamma_{s(x)}$. Moreover, f is measurable: If $B \in \mathfrak{B}$ is a Borel set, then as noted above, $f^{-1}(B) \in \mathfrak{A}'$. If $B' \in \mathfrak{B}' \setminus \mathfrak{B}$, then $B' \subseteq B$, where B is a Borel set with $\nu(B) = 0$. But this implies that $\mu(f^{-1}(B)) = 0$. Since $f^{-1}(B') \subseteq f^{-1}(B)$, we find that $f^{-1}(B') \in \mathfrak{A}'$. The exact same argument shows f^{-1} to be measurable as well, which establishes the claim. \square

Our goal is to approximate a given dynamical system (X, μ, S) with a finite system. Accordingly, suppose that X' is a graph whose vertices are labeled with elements of X (i.e. X' is an X -labeled set) and whose edges are directed and labeled with elements of S . If μ' is a probability measure on X' , then we claim that μ' naturally determines a probability measure ν' on $\text{Sch}(S, X)$: Given an r -neighborhood $U \in \text{Sch}_r(S, X)$, set

$$\nu'(C_U) = \sum_{B_r(x) \cong U} \mu'(x)$$

as long as there exists an $x \in X'$ such that $B_r(x) \cong U$ when the vertex-labels of X' are truncated to their first r digits. Once this requirement is met, ν' may be extended arbitrarily to other cylinder sets in such a way that it becomes a probability measure (the measure μ' therefore technically defines a family of measures on $\text{Sch}(S, X)$, but the choice of representative will not matter for us). Note that if $K \subseteq S$ and $\psi : K \rightarrow \text{End}(X')$ is a (K, ε) -action on an X -labeled set X' , then X' comes equipped with a natural graph structure, which consists of all triples $(x, s, \psi_s(x))$, where $x \in X'$ and $s \in K$.

Definition 4.4. Let (X, μ, S) be a dynamical system, let $K \subseteq S$, and let $\varepsilon > 0$. A triple (X', μ', ψ) , where X' is a finite X -labeled set, μ' is a probability measure on X' , and $\psi : K \rightarrow \text{End}(X')$ is a (K, ε) -action, is a (K, ε) -approximation to (X, μ, S) if

$$\|\mu' - \psi(s)_*\mu'\| \leq \varepsilon$$

for all $s \in K$ and

$$\left| \int f d\nu - \int f d\nu' \right| \leq \varepsilon$$

for any bounded continuous function $f : \text{Sch}(S, X) \rightarrow \mathbb{R}$, where ν and ν' are the probability measures on $\text{Sch}(S, X)$ determined by (X, μ, S) and (X', μ', ψ) , respectively.

We thus require the measure μ' on X' to be approximately invariant with respect to ψ and the measure ν' to be close, in the weak topology, to the S -invariant measure ν .

Definition 4.5. A dynamical system (X, μ, S) is *sofic* if for any finite $K \subseteq S$ and any $\varepsilon > 0$, it admits a (K, ε) -approximation.

It is worth pausing to point out that our approach mimics, in many ways, the approach of Elek and Lippner, which in turn mimics the idea behind so-called Benjamini-Schramm convergence, introduced in [2]. Here a sequence of graphs (or other objects—in particular, graphs which carry additional structure) is interpreted as a sequence of probability measures on the space of rooted graphs by choosing the position of the root of each graph in the sequence uniformly at random. Each such measure is an invariant (or *unimodular*) measure, and one may study the corresponding weak limit of the sequence. The key difference with our approach is that we consider general (approximately) invariant measures on finite structures, and not merely those which arise upon choosing the root uniformly at random. We conclude with a basic example of a sofic dynamical system.

Example 4.6. Denote by $\mathbb{S}^1 = [0, 1]/\sim$ the circle, obtained by gluing together the endpoints of the unit interval, and consider the classical angle-doubling map $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ given by

$$f(x) = 2x \pmod{1}.$$

The map f is 2-to-1, and the Lebesgue measure λ on \mathbb{S}^1 inherited from $[0, 1]$ is invariant with respect to the \mathbb{N} -action determined by f . An element of $(\text{Sch}(\mathbb{N}, X), \nu)$, the corresponding space of full Schreier graphs, is almost surely a rooted tree each of whose vertices has two incoming f -labeled edges attached to it and one outgoing f -labeled edge attached to it (here we have made an \mathbb{N} -equivariant identification between \mathbb{S}^1 and $X = \{0, 1\}^\infty$ and, abusing notation, will also denote by λ the image of the Lebesgue measure under this identification). To construct a sofic approximation to the system $(\mathbb{S}^1, \lambda, \mathbb{N})$, let X' be the disjoint union of all possible preimage trees of elements $\omega \in \{0, 1\}^\infty$, up to depth k . That is, a connected component of X' consists of an X -labeled vertex at level 0, together with two X -labeled preimage vertices at level 1, and so on up to level k , which consists of 2^k X -labeled preimage vertices. A connected component U of X' also determines a cylinder set C_U , which has a certain measure. Define a measure on X' by assigning mass $2^{-k}\lambda(C_U)$ to each vertex at level k of a connected component U of X' , then normalize to obtain a probability measure λ' .

It is not difficult to see that, provided k is chosen to be sufficiently large, the measure λ' is approximately invariant with respect to the elements

$$\{f, \dots, f^{o_n}\} \subseteq \mathbb{N}$$

and moreover that the associated measure ν' on $\text{Sch}(\mathbb{N}, X)$ approximates ν in the weak topology. Choosing ever larger values of k and carrying out this construction shows the system $(\mathbb{S}^1, \lambda, \mathbb{N})$ to be sofic.

It would appear that there is much to investigate concerning our definition of a sofic dynamical system. Idle questions include: Is every \mathbb{N} -action sofic? (We conjecture a positive answer.) Is every dynamical system (in our sense) sofic? Is it possible to work only with the full orbit equivalence relation determined by a semigroup action, thereby developing a theory analogous to that of Feldman and Moore (see [10])? Suppose a dynamical system (X, μ, S) admits a *natural extension* $(\tilde{X}, \tilde{\mu}, S)$ that makes each $s \in S$ invertible—what, if any, is the relationship between soficity of the original system and soficity of the extension? Which semigroups admit faithful actions by measure-preserving endomorphisms on a Lebesgue space?

A further idea is implicit in our work. Having relaxed the requirement that sofic approximations to a dynamical system come equipped with a uniform measure, it is natural to ask whether there is a reasonable way to approximate dynamical systems which come only with a quasi-invariant measure (so that the measure class, rather than the measure itself, is preserved), or even an arbitrary measure, perhaps by requiring the Radon-Nikodym derivatives associated to the finite approximating system to converge to those of the original system. It could be interesting to develop soficity in this direction as well and examine its relationship with soficity for dynamical systems with an invariant measure. We may pursue this idea in a future work.

REFERENCES

- [1] D. Aldous and R. Lyons, *Processes on unimodular random networks*, Electron. J. Probab. **12** (54) (2007), p. 1454–1508.
- [2] I. Benjamini and O. Schramm, *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (23) (2001), p. 1–13.
- [3] L. Bowen, *Measure conjugacy invariants for actions of countable sofic groups*, J. Amer. Math. Soc. **23** (1) (2010), p. 217–245.
- [4] L. Bowen, *Entropy theory for sofic groupoids I: The foundations*, J. Anal. Math. **124** (1) (2014), p. 149–233.
- [5] J. Cannizzo, *The boundary action of a sofic random subgroup of the free group*, To appear in Groups Geom. Dyn.
- [6] T. Ceccherini-Silberstein and M. Coornaert, *On sofic monoids*, Semigroup Forum **89** (3) (2014), p. 546–570.
- [7] K. Dykema, D. Kerr, and M. Pichot, *Sofic dimension for discrete measured groupoids*, Trans. Amer. Math. Soc. **366** (2) (2014), p. 707–748.

- [8] G. Elek and G. Lippner, *Sofic equivalence relations*, J. Funct. Anal. **258** (2010), p. 1692–1708.
- [9] G. Elek and E. Szabó, *Hyperlinearity, essentially free actions and L^2 -invariants: The sofic property*, Math. Annal. **332** (2005), p. 421–441.
- [10] J. Feldman and C. C. Moore, *Ergodic equivalence relations, cohomology, and von Neumann algebras I*, Trans. Amer. Math. Soc. **234** (2) (1977), p. 289–324.
- [11] M. Gromov, *Endomorphisms of symbolic algebraic varieties*, J. Eur. Math. Soc. **1** (2) (1999), p. 109–197.
- [12] M. Kambites, *A large class of sofic monoids*, Preprint (2014).
- [13] V. G. Pestov, *Hyperlinear and sofic groups: A brief guide*, Bull. Symb. Log. **14** (4) (2008), p. 449–480.
- [14] B. Weiss, *Sofic groups and dynamical systems*, Sankhya A **62** (3) (2000), p. 350–359.

DEPARTMENT OF MATHEMATICAL SCIENCES
STEVENS INSTITUTE OF TECHNOLOGY
1 CASTLE POINT ON HUDSON
HOBOKEN, NJ 07030

E-mail address: `jan.c.cannizzo@gmail.com`